Improving the Instance Selection Method for Better Detection of Depression in Children and Adolescents

Ariane C. B. da Silva[1][a], Maycoln L. M. Teodoro[2][b] and Cristiane N. Nobre[1][c]

[1] *Institute of Exact Sciences and Informatics, Pontifical Catholic University of Minas Gerais, Dom José Gaspar, Belo Horizonte, Brazil*

[2] *Department of Psychology, Federal University of Minas Gerais, Belo Horizonte, Brazil*

Keywords: Depression, Adolescence, Children, Machine Learning, Instance Selection.

Abstract: Depression is the leading global cause of disability and often begins in adolescence, a critical period for developing depressive symptoms. Major depressive disorder in the early stages of life is common worldwide but challenging to diagnose. Identifying the most striking profiles of depression in children and adolescents could benefit the training and performance of Machine Learning models and thus help in the diagnosis. Instance Selection is one of the most applied methods for data reduction, allowing the most significant samples to represent them. This work seeks to improve the SI with the Ant Colony Optimization heuristic, introducing stochasticity control to better characterize profiles of children and adolescents with depression. The proposed technique increased the detection rate of individuals with high symptoms in all evaluated algorithms between 0.07 and 8.93 percentage points.

1 INTRODUCTION

Depression is a leading cause of disability around the world and contributes significantly to the global burden of disease. The World Health Organization (WHO, 2022) estimates that over 300 million people live with depression worldwide. It is the most significant contributor to deaths by suicide (almost 800 thousand per year) and the most critical contributor to global disability (7.5% of all years lived with disability). Despite being common worldwide, the diagnosis of depression in adolescence is still challenging since it presents a wide range of symptoms that can be confused with the natural alterations pertinent to this period of life. In addition, Johnson et al. (2018) relate depression in adulthood to its onset in the early stages of life and emphasize the importance of identifying it and starting treatment as soon as possible.

In Machine Learning (ML), the performance of classification algorithms depends on the training data's quality. Thus, removing noise, outliers, and other instances from the training set that could be harmful or misleading for the algorithm that learns a model is crucial. One widely applied method is Instance Selection (IS), whose main objective is to select the most significant instances of the original base. The IS issue represents a combinatorial optimization task that several heuristics can solve (Salama et al., 2016). This work used the Ant Colony Optimization (ACO) heuristic (Dorigo et al., 2006) due to its characteristics of being able to be applied to different discrete optimization problems with relatively few modifications (essential to generalize the possibilities of use in different contexts of the library that we made available), can be used in dynamic applications, is little affected by the initialization condition and is less likely to get stuck in local optima than conventional greedy algorithms (Salama et al., 2016).

Concerning research involving Data Reduction with IS and ACO, much scientific effort has been employed to look for a reduced set of instances to mitigate the low computational efficiency and high storage requirements (Salama et al., 2016), (Miloud-Aouidate and Baba-Ali, 2013), (Akinyelu, 2020), (Gong et al., 2021), (El Bakrawy et al., 2022), (Hott. et al., 2022). However, in the context of health, IS with ACO can be used with a more specific objective. According to Salama et al. (2016), IS is beneficial to reduce the training time and improve the characterization of the instances, which would be of great value

[a] https://orcid.org/0000-0003-2477-4433
[b] https://orcid.org/0000-0002-3021-8567
[c] https://orcid.org/0000-0001-8517-9852

B. da Silva, A., Teodoro, M. and Nobre, C.
Improving the Instance Selection Method for Better Detection of Depression in Children and Adolescents.
DOI: 10.5220/0012355600003657

In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2024) - Volume 2*, pages 404-411
ISBN: 978-989-758-688-0; ISSN: 2184-4305

in applications with health-oriented databases, such as the investigation of depression, providing a more representative model. Unlike the typical approach in works related to IS, in this paper, the selected instances were evaluated for the performance achieved in the classifiers and qualitatively to characterize their depression profiles better.

Thus, this study aims to apply the IS with ACO to obtain more efficient classification models to identify depression in children and adolescents and seek a better characterization of their profiles. Besides, we introduced to the algorithm a new parameter to control the probability of instance selection, allowing us to adjust the algorithm for a more or less exploratory search in the sample space. Finally, a Python library package was available[1] for using the algorithm in other contexts.

2 BACKGROUND

2.1 Instance Selection with ACO

The Instance Selection (IS) technique requires a complete search of all possible combinations of instances to find the best set. This study used the ACO heuristic (Dorigo et al., 2006) to find the appropriate subset of data that best characterizes the original set based on the accuracy obtained by the classifier *k*NN. ACO is a stochastic search method inspired by the natural behavior of ants, which seek the shortest path between the nest and food, depositing pheromones in the soil to mark the path other colony members must follow. ACO exploits a similar mechanism to solve optimization problems.

The problem search space is represented as a graph: the input instances represent the vertices, and the Euclidean distance between them gives the edges. Each artificial ant starts from a different instance and navigates the graph, creating subsets that are submitted to an ML algorithm to evaluate them according to the achieved performance. The probability of an ant selecting an instance is based on the heuristic advantage associated with the instance and the amount of pheromone present Salama et al. (2016). The algorithm returns the best set as the response according to the best accuracy. Algorithm 1 presents a basic pseudocode of the IS with ACO.

This paper implemented the IS algorithm based on the ACO principles proposed in (Miloud-Aouidate and Baba-Ali, 2013), called ANT-IS. It was chosen

Data: Any combinatorial problem
Result: Best solution
initialization;
while *stop condition not reached* **do**
    AntsBuildSolutions;
    UpdatePheromone;
end
EvaluateSolutions;

Algorithm 1: Basic ACO pseudocode.

because it presents a simple and versatile IS approach that can be easily adapted to perform attribute selection and allows execution in parallel using several processing cores. That article describes the main steps and equations that indicate ant behavior and the calculations of an instance's heuristic advantage[2].

3 RELATED WORKS

Regarding IS with ACO, works generally aim to reduce the computational time for training ML models Salama et al. (2016), Akinyelu (2020), Gong et al. (2021), El Bakrawy et al. (2022). Salama et al. (2016) presented five improved versions of their own ACO-based IS algorithm, incorporating feature selection. Akinyelu (2020) applied a threshold detection approach and IS technique (ACO + *k*NN) to improve the speed of big data classification models.

Although the articles mentioned above have achieved excellent data reduction results, none of the previous approaches can change the reduction percentage. In healthcare databases, generally small and very unbalanced, adjusting the selection probability of an instance allows control of the amount of reduction to which a dataset will be subjected, and, in this way, we can obtain more customized response sets.

Hott. et al. (2022), in turn, applied IS with ACO to obtain more efficient classification models in identifying school performance in arithmetic, reading, and writing of children and adolescents with hyperactivity disorder and attention deficit.

As in the previous study, this work also seeks to apply the IS and ACO techniques in the health context, specifically depression. However, in our article, the selected instances will be evaluated visually in the sample space and also regarding their representative quality within their class. In this context, we are not reducing the data due to the size of the used database, , which is not big, the goal here is to reduce it in order to find a subset of the original database that allows the

[1] Installation and use instructions can be found at https://test.pypi.org/project/antcolony-is/

[2] A simplified simulation can be seen here https://youtu.be/kdO1-36rvok

training of ML algorithms with a better detection rate of depression in children and teenagers.

4 MATERIALS AND METHODS

4.1 Dataset

The database used in this study was obtained in partnership with the Graduate Program in Psychology: Cognition and Behavior at the Federal University of Minas Gerais/Brazil (UFMG). The dataset[3] contains 377 instances and 75 attributes, with information on children and adolescents (10 - 16 years old) with different depressive symptoms. In order to adapt the raw data, it was necessary to submit them to the following pre-processing procedures.

1. *Identification and manipulation of the class attribute*. As the database originally received was not classified, responses to the Childhood Depression Inventory (CDI) questionnaire were used for this purpose. Item scores are summed into a total depression score (CDI Sum), which ranges from 0 to 54. The higher the score, the greater the chances the patient has a higher depressive state (Bang et al., 2015). However, the CDI score alone does not determine the existence or not of depression, but evidence that supports the assessment made by the professional. At this stage, an instance that did not have CDI information had to be removed. The literature has no unanimity regarding the cutoff value determining the division between high and low symptomatology. In this study, Kovacs and Staff (2003)'s recommendation was considered, regarding using the 85th percentile to indicate high depressive symptomatology.
2. *Database balancing*. "LOW" and "HIGH" symptomatology classes have 314 and 63 individuals, respectively. Such an imbalance could interfere with the proposed instance selection process, tending to obtain better results for the majority class to the detriment of the minority, the main target of this study. Therefore, balancing techniques such as oversampling and random subsampling were performed to compare the performance obtained by IS with and without prior balancing.
3. *Training and test set splitting*. The database was divided into two groups, one for model training and validation, and another for testing, for each of the proposed balancing scenarios. The divisions performed can be viewed in Table 1.

[3]A description of all database attributes used as predictors to the ML models can be found on https://docs.google.com/spreadsheets/d/15rsErDkY3xFCG3Rubl1QoZeT8GBMQaXUPgPKWZHmgjY/edit?usp=sharing

Table 1: Number of instances per class for training/validation and testing.

Class	Original dataset	Unbalanced train/val	Under train/val	Over train/val	Test
HIGH	63	48	48	243	15
LOW	314	243	48	243	80
Total	**377**	**291**	**126**	**486**	**86**

4.2 IS + ACO, Balancing Methods, and ML Algorithms

ANT-IS algorithm was implemented in Python v3.10 and the tests were carried out using InstanceSelection (*antcolony-is* package v1.0.1) from the *TestPyPI* index, with default settings. The p parameter was added to the algorithm, which controls the degree of stochasticity, allowing adjustment of instance selection probability. For p = 1 the selection probability is maximum. For p close to zero the probability is minimum, directly influencing the size of the reduced set provided in the ANT-IS output. The experiments used values of p = 0.3, 0.5 and 0.7.

Regarding the balancing of the original dataset, two techniques were applied. Undersampling was done randomly, selecting from the majority class the same number of instances in the minority class, and Oversampling using SMOTE (*over_sampling* package) from the *imblearn* library package v0.9.1, with Python v3 default settings. Table 1 describes the final proportions obtained after each of these steps.

As for the ML algorithms, after the IS performed by the Ant Colony, the obtained response set was provided for the training of five classification algorithms to evaluate the performance of the Ant-IS: 1NN, CART, neural network MLP, SVM and RF. All of them were built using the *Scikit-learn* library package version 1.0.2, with default settings. The experiments were performed on Windows 11 operating system using an Intel(R) Core(TM) i7 processor, 2.60GHz, 16GB of RAM and the PyCharm v2022.1.3 tool. Figure 1 outlines the used methodology.

4.3 Model Quality Assessment Metrics

Precision[4], *Recall*[5], and *F-measure*[6] metrics were used to assess the quality of the ML models. Preci-

[4]$Precision = \frac{VP}{VP+FP}$

[5]$Recall = \frac{VP}{VP+FN}$

[6]$F-Measure = \frac{2\times Recall\times Precision}{Recall+Precision}$

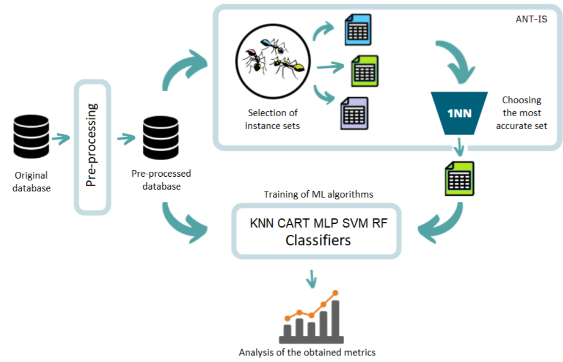


Figure 1: Selection proceeding and results evaluation.

sion is the rate of instances correctly classified as belonging to the class in question out of all those classified in the class. *Recall* refers to the percentage of class instances that were correctly predicted to belong to the class. The *F-measure* is a harmonic mean between Precision and *Recall*. The training of the ML models was carried out through a stratified 10-fold cross-validation method, in which the train-validation procedure is repeated ten times and the mean value represents the test result.

5 RESULTS AND DISCUSSION

We analyzed three balancing test scenarios. 1) IS on the original unbalanced data, 2) IS over oversampled data (SMOTE), 3) IS over subsampled data (random subsampling technique). For this balancing test step, the selection probability was set to 50% ($p = 0.5$). Figure 2 presents the results of the tests carried out. Regarding the influence of data balancing before instance selection, it is interesting that the ANT-IS algorithm performed well on unbalanced data. It outperformed the other balancing techniques in 7 of the 10 possibilities in the F-Measure metric, even not being a balancing algorithm itself. In the Precision and Recall metrics, the technique that obtained the best performance was subsampling, corroborating the idea of the positive influence of balanced data in the training of ML models. Therefore, the other experiments in this study were conducted on the subsampled database and the IS performed on it.

The average reduction rate obtained with the parameter p set to 0.5 (50% probability of selecting an instance) was 48%, and Table 2 gathers the metrics evaluated in this condition. Table 2 summarizes the gain or loss obtained, in percentage points, using the reduced set given as output from the ANT-IS, compared to the subsampled complete set, for each classifier. There was a significant gain in some specific cases, in others, a considerable reduction, and, in others, the values remained close, with variations around 0 to 2 percentage points, more or less. Negative values indicate that there was a reduction in the value of the evaluated metric. The most impacted algorithm by the IS was the MLP neural network, in the *Precision* metric, for the HIGH symptomatology class, reducing 5.59 percentage points. Neural networks need a significant amount of data for their learning, which could explain the low performance, in this metric concerning the other classifiers. However, the *Recall* metric rose 8.93 percentage points for the same class and classifier, providing a harmonic average *F-Measure* with a gain of 7.05 percentage points. Although the results oscillate between good gains and slight reductions, the objective of the work to better identify depression profiles suggests a more detailed analysis of the *Recall* metric, which is also considered a detection rate. Based on the class of HIGH symptomatology, the main target of this study, all classifiers obtained a gain in that metric when using the reduced training set. Such results indicate that the technique used tends to better detect these individuals than the results obtained without using it. Regarding the performance and scalability of ANT-IS, the size of the database strongly impacts the method's execution time, as illustrated in Figure 3. The diameter of the circles represents the number of attributes in each database.

Table 2: Gain/Reduction obtained with ANT-IS (in percentage points).

	Class	Precision	Recall	F-Measure
1NN	HIGH	4,40	5,66	4,69
	LOW	2,14	2,38	2,06
CART	HIGH	0,67	4,07	1,14
	LOW	1,11	-2,00	-1,34
MLP	HIGH	-5,59	8,93	7,05
	LOW	1,40	-0,21	3,43
SVM	HIGH	-0,51	0,07	-0,59
	LOW	-0,11	-1,00	-0,76
RF	HIGH	-0,55	0,10	-0,58
	LOW	-0,07	-0,98	-0,73

Incorporating the p parameter to ANT-IS allows us to further improve the metrics of each of the ML algorithms according to the problem's context. If a more significant reduction of data in the IS is needed, lower p values can be experimented with. If the objective is to increase the quality of the ML model, it is possible to reach better metrics with higher values of p. At this stage, three scenarios were evaluated, adjusting the p probability of selecting an instance set to i) 30%, ii) 50% and iii) 70%. It is worth remembering that with the increase in the probability of selecting an instance, there is also an increase in the

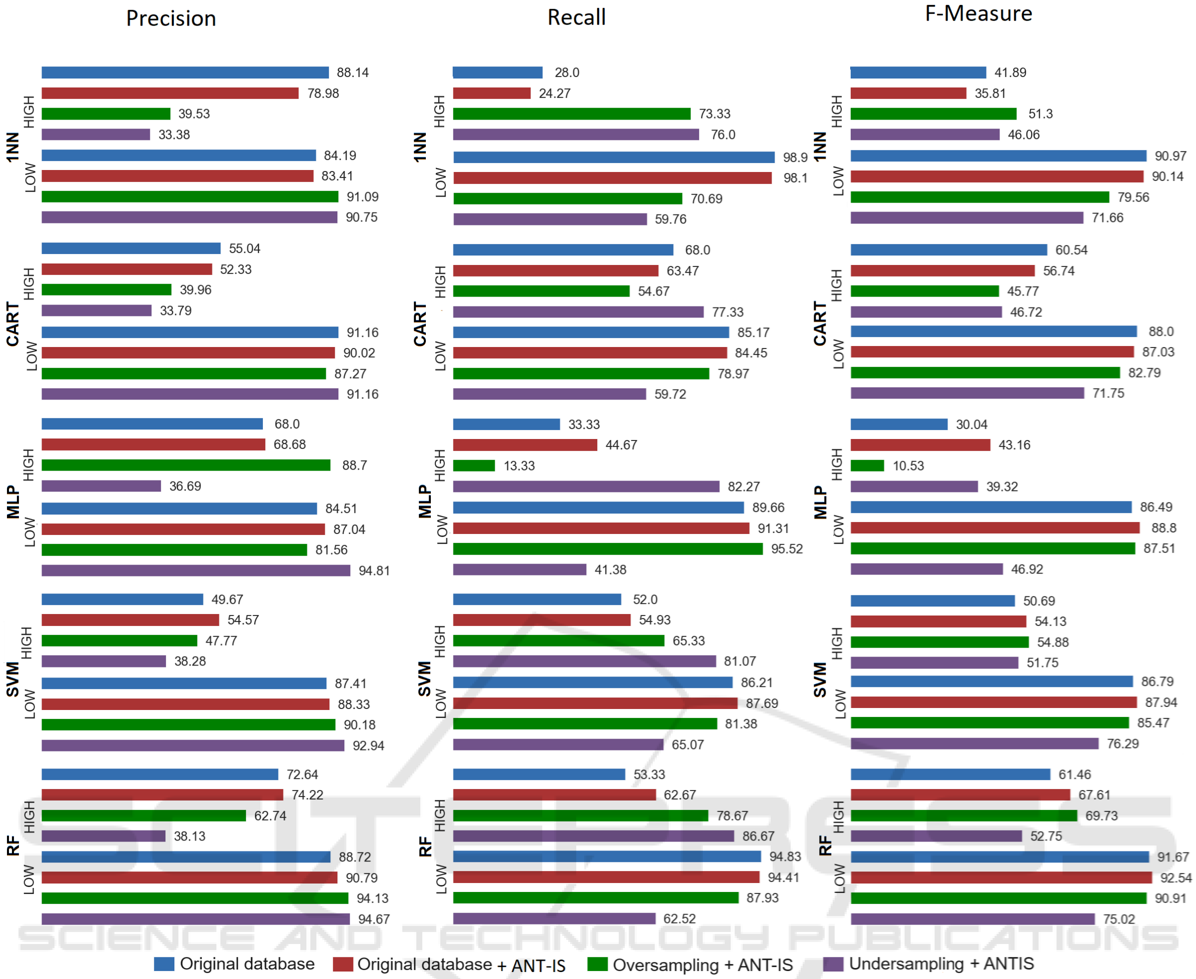


Figure 2: Results of each experiment on the unbalanced database.

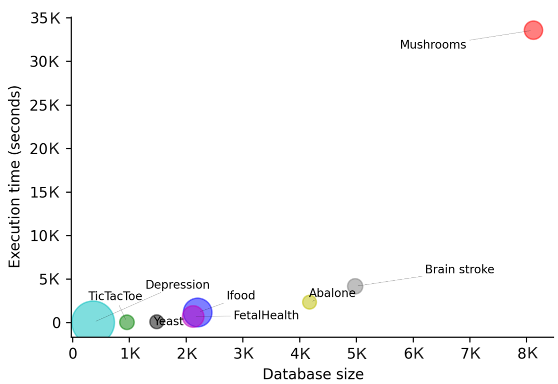


Figure 3: Scalability of the ANT-IS algorithm.

number of instances in the response set of the ANT-IS algorithm, and consequently, the ML model will be trained with a more extensive training set. The experiments carried out consisted of running the ANT-IS on the database with each of the mentioned probabilities and observing the value of the metrics, comparing them with those obtained when using the subsampled complete set in the training of the ML algorithm.

Figure 4 represents the impact of varying the probability of selecting an instance on the classifiers' metrics. For the evaluated database, the KNN, CART and RF algorithms showed a proportional improvement in the metrics with the increase in the reduced training set. In contrast, the SVM algorithm showed the opposite behavior, reducing the gain in the metrics with the increase in the training set. The MLP neural network, on the other hand, presented a different and atypical behavior concerning the others. The objective of this experiments stage was not to prove which algorithm obtained the best performance (which would require a careful adjustment of its hyperparameters) but rather to identify that they are strongly impacted by the size of the reduced set, demonstrating the importance of having a way to adjust the instance selection probability.

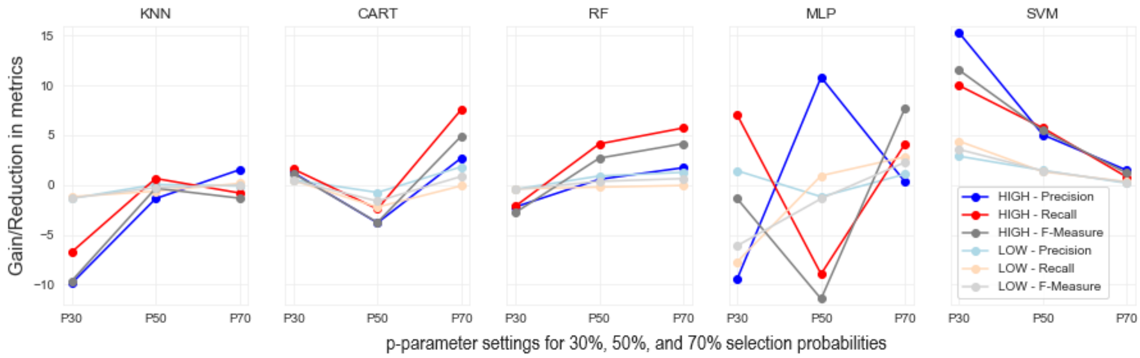


Figure 4: Gain/Reduction in metrics due to p variation.

Another important point to be considered is the identification of the most repeatedly selected instances by the Ant-IS algorithm since, due to its stochastic characteristic, it presents a different set of instances as a result of each execution. Figure 5 and Figure 6 illustrate the selection frequency of each instance in 100 repeated executions of the algorithm on the analyzed database.

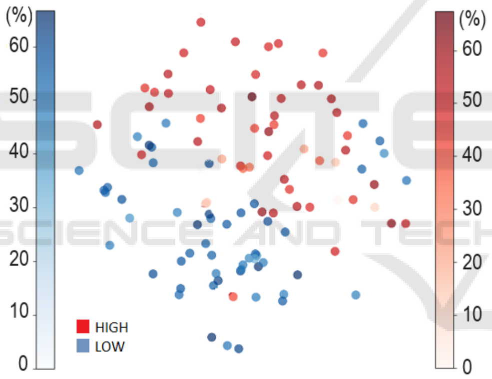


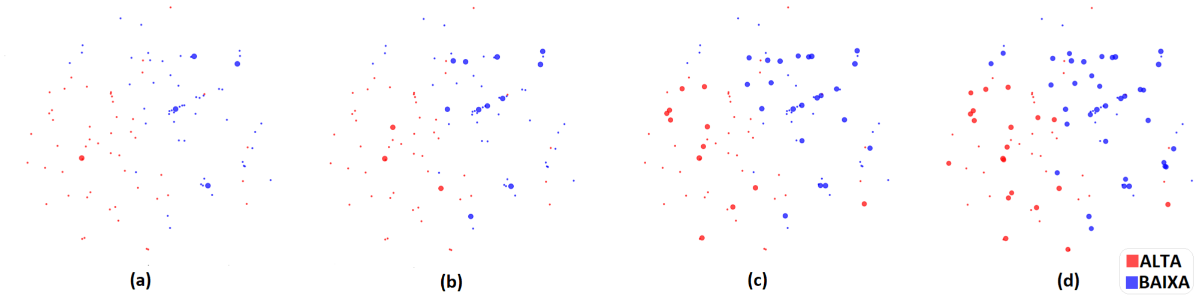
Figure 5: Frequency of instance selection by Ant-IS in 100 runs.

In Figure 5, each point identifies an instance duly represented in the sample space, associated with its symptom classification. The selection covered the sample space well, not concentrating on any specific area. This feature of ANT-IS avoids getting stuck in local minima in the search space. It is also noted, through Figure 6, that there was a slight tendency for Ant-IS to select instances of the LOW symptomatology class. Of the 57 instances selected more than half the time, 36 belonged to the LOW symptomatology class and 21 to the HIGH one.

Moving on to a more specific analysis, based on the 3 instances that were most selected in their respective classes, Table 3 describes their most relevant attributes. There are three male and three female individuals, aged between 11 and 15 years old. Three are LOW class and three are HIGH class. Three were receiving regular psychological care, two of them with HIGH symptomatology. Only one of the instances has the parents split up, and it is a LOW symptomatology instance. The combination of attributes of HIGH-class instances presents medium-high values concerning others, for *anxiety*, *social problems*, and *conduct*, in addition to *negative or self-defeating thoughts*. Other striking features of instances of HIGH symptomatology are the presence of *oppositional defiant disorder*, *high aggressiveness*, *difficulty paying attention* and *externalizing disorders*. Such conditions were observed by a psychology professional at the time of collecting information from children and adolescents in this study, based on the Diagnostic and Statistical Manual of Mental Disorders V (APA et al., 2014) and the applied CDI and YSR questionnaires. The last line of Table 3 presents the minimum and maximum values existing in the evaluated database, to help the comparative analysis.

About the instances of the LOW symptomatology class, except for some specific attributes with higher values (*negative thinking* and *anxiety* with high values in one of them), the other attributes remained close to the lower limits of the analyzed sample. However, such symptoms may represent warning signs for this individual in question. Regarding the time spent with the parents, in all six instances, the number of hours spent with the mothers was slightly higher than those spent with the fathers, or the same in a few cases. All of them reported that the father worked outside the home, and only two reported that the mother did not work, one from the LOW class and the other from the HIGH symptomatology class.

The results of this study corroborate others in the literature in the sense that depression can be associated with other psychiatric disorders and comorbidities. According to Maughan et al. (2013), two-thirds of adolescents with depression have at least one comorbid psychiatric disorder, and 10-15% have two



(a) TOP5 most selected instances (appeared in 65 out of 100 ANT-IS runs); (b) TOP15 (were selected in 60 out of 100 executions); (c) TOP30 (appeared in 55% of runs) and (d) instances that were selected in 50% of ANT-IS runs.

Figure 6: Separation by selection frequency bands.

Table 3: Relevant attributes of the 3 most selected instances in each class.

Class	BR	ED	Aggr	OD	CP	Age	S	PS	SP	Anx	Withdr	NT	AD	SC	SF
LOW	50	47	52	51	50	15	M	No	54	69	54	77	50	67	67
LOW	50	40	50	50	50	11	F	No	51	54	55	50	51	60	67
LOW	50	44	50	51	50	14	F	Yes	58	56	50	50	60	68	66
HIGH	57	72	87	76	67	12	F	Yes	58	66	55	73	73	68	67
HIGH	58	62	66	61	62	14	M	No	75	74	60	55	54	58	61
HIGH	60	66	68	61	64	13	M	Yes	63	70	57	60	73	64	61
Min-Max	50-79	29-80	50-89	50-80	50-78	10-16	-	-	50-90	50-100	50-100	50-95	50-100	50-90	0-100%

BR = Break Rules. ED = Externalizing Disorders. Aggr = Aggressiveness. OD = Oppositional defiant disorder.
CP = Conduct problems. S = Sex. PS = Psychological support. SP = Social problems. Anx = Anxiety. Withdr = Withdrawal.
NT = Negative thinking. AD = Attention difficulty. SC = Somatic complaints. SF = Selection frequency.

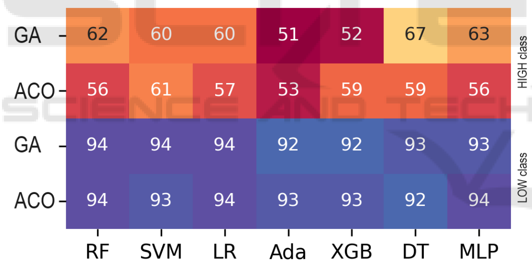


Figure 7: Comparison between ACO and Genetic Algorithm (GA) approaches for IS according to the F-Measure metric.

or more associated comorbidities. Adolescents with depression are more likely to have anxiety and more likely to also have a disruptive behavior disorder compared to those who are not depressed. The three most selected instances of HIGH symptomatology showed this picture of associated disorders.

Finally, regarding comparing the ACO with other approaches for IS, we compared it with the Genetic Algorithm (GA) heuristic (Santana, 2021) in this article. Figure 7 shows that, for the depression database analyzed in seven ML algorithms (RF, SVM, Logistic Regression, Adaboost, XGBoost, Decision Tree, and MLP), the F-Measure metric remained very close in both heuristics, with the ACO heuristic taking only 1/6 of the time spent by the GA for IS on average, presenting practically the same reduction rate (48% and 49% for ACO e GA, respectively).

6 CONCLUSIONS

With regard to obtaining more assertive classification models, the technique employed proved to be satisfactory, given the average reduction of 48% in the size of the original data and the increase in the *Recall* detection rate (between 0.07 and 8.93 percentage points depending on the ML algorithm evaluated) for the HIGH symptomatology class, the main target of this study. This fact could indicate that the selected instances would be the ones that best characterize the symptomatology of depressive disorder in children and adolescents for the analyzed database, in terms of classification. However, the experiments showed a tendency of the algorithm to select more LOW symptomatology instances, even with the training data balancing. Thus, the technique still needs further improvement, more tests and support from the analysis of a psychology professional in evaluating the results obtained.

The use of only one specific database makes it impossible, in principle, to generalize the results achieved by the algorithm to other contexts. Another point of attention is the more careful adjustments of the employed ML algorithms' parameters that were instantiated with their general typical values. Better adjustments could achieve different results. It is also worth mentioning the small size of the base evaluated, which can directly influence the quality of the models generated and the results achieved.

As for future work, three main points need to be worked on: optimizing the algorithm's performance, since preliminary tests on larger databases proved to be still too slow; expanding the number of databases tested, including others of different sizes, both balanced and unbalanced, to investigate the balancing capacity of ANT-IS better and generalize its use; and finally validate the attribute selection introduced to the instance selection algorithm, evaluating whether its application produces any improvement in classification metrics, favoring its application in big data contexts.

ACKNOWLEDGEMENTS

The authors thank the National Council for Scientific and Technological Development of Brazil (CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico – Code: 311573/2022-3), the Pontifícia Universidade Católica de Minas Gerais – PUC-Minas, the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES – Grant PROAP 88887.842889/2023-00 – PUC/MG, Grant PDPG 88887.708960/2022-00 – PUC/MG - Informática and Finance Code 001), and the Foundation for Research Support of Minas Gerais State (FAPEMIG – Code: APQ-03076-18).

REFERENCES

Akinyelu, A. A. (2020). Bio-inspired technique for improving machine learning speed and big data processing. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

APA et al. (2014). *DSM-5: Manual diagnóstico e estatístico de transtornos mentais*. Artmed Editora.

Bang, Y. R., Park, J. H., and Kim, S. H. (2015). Cut-off scores of the children's depression inventory for screening and rating severity in korean adolescents. *Psychiatry investigation*, 12(1):23.

Dorigo, M., Birattari, M., and Stutzle, T. (2006). Ant colony optimization. *IEEE computational intelligence magazine*, 1(4):28–39.

El Bakrawy, L. M., Cifci, M. A., Kausar, S., Hussain, S., Islam, M. A., Alatas, B., and Desuky, A. S. (2022). A modified ant lion optimization method and its application for instance reduction problem in balanced and imbalanced data. *Axioms*, 11(3):95.

Gong, C., Su, Z.-g., Wang, P.-h., Wang, Q., and You, Y. (2021). Evidential instance selection for k-nearest neighbor classification of big data. *International Journal of Approximate Reasoning*, 138:123–144.

Hott., H., Jandre., C., Xavier., P., Miloud-Aouidate., A., Miranda., D., Song., M., Zárate., L., and Nobre., C. (2022). Selection of representative instances using ant colony: A case study in a database of children and adolescents with attention-deficit/hyperactivity disorder. In *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies - HEALTHINF,*, pages 103–110. INSTICC, SciTePress.

Johnson, D., Dupuis, G., Piche, J., Clayborne, Z., and Colman, I. (2018). Adult mental health outcomes of adolescent depression: A review. *Depression and anxiety*, 35(8):700–716.

Kovacs, M. and Staff, M. (2003). Children's depression inventory (cdi): Technical manual update. multi-health systems. *Inc.: North Tonawanda, NY, USA*.

Maughan, B., Collishaw, S., and Stringaris, A. (2013). Depression in childhood and adolescence. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 22(1):35.

Miloud-Aouidate, A. and Baba-Ali, A. R. (2013). An efficient ant colony instance selection algorithm for knn classification. *International Journal of Applied Metaheuristic Computing (IJAMC)*, 4(3):47–64.

Salama, K. M., Abdelbar, A. M., and Anwar, I. M. (2016). Data reduction for classification with ant colony algorithms. *Intelligent Data Analysis*, 20(5):1021–1059.

Santana, R. C. (2021). *Algoritmo genético de aprendizado ativo para bases de dados pequenas e desbalanceadas no contexto de depressão em crianças e adolescentes*. PhD thesis, (texto da Qualificação). Pontifícia Universidade Católica de Minas Gerais. Programa de Pós Graduação em Informática.

WHO (2022). *World mental health report: transforming mental health for all*. World Health Organization.